UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/711,114	08/24/2004	Christopher S. Jochumson	22207-010120	5113

51111 7590 06/13/2006

AKA CHAN LLP
900 LAFAYETTE STREET
SUITE 710
SANTA CLARA, CA 95050

EXAMINER
LERNER, MARTIN

ART UNIT	PAPER NUMBER
2626	

DATE MAILED: 06/13/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/711,114	JOCHUMSON, CHRISTOPHER S.	
	Examiner	Art Unit	
	Martin Lerner	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *09 May 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 to 3, 5 to 6, and 8 to 21* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1 to 3, 5 to 6, and 8 to 21* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Election/Restrictions

Applicant's election without traverse of Group I, Claims 1 to 3, 5 to 6, and 8 to 21 in the reply filed on 09 May 2006 is acknowledged.

Claims 4 and 7 are withdrawn from further consideration pursuant to 37 CFR 1.142(b) as being drawn to a nonelected invention, there being no allowable generic or linking claim. Election was made **without** traverse in the reply filed on 09 May 2006.

Double Patenting

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. See *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent is shown to be commonly owned with this application. See 37 CFR 1.130(b).

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

Claims 1 to 3, 5 to 6, and 9 to 21 are rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 1 to 26 of U.S. Patent No. 6,865,536 in view of *Meisel et al*.

Although the conflicting claims are not identical, they are not patentably distinct from each other because the current claims of the application and the prior claims of the patent set forth the same subject matter of two or more clients storing audio speech in one or more buffers and a server comprising the capability to receive packets from each of the at least two clients. The only significant feature omitted by the claims of the parent patent is storing audio speech in buffers in a raw uncompressed audio format, as the claims of the parent patent do not expressly say that buffers store raw uncompressed speech. However, *Meisel et al.* teaches preprocessing for speech recognition, where a general approach is disclosed of a means for buffering raw analog or digitized speech data for analysis by collecting and storing the raw data. Optimal parameters can then be extracted by analysis. (Column 5, Lines 6 to 12) It would have been obvious to one having ordinary skill in the art to modify the claims of the parent patent to include a feature of storing raw uncompressed audio speech in buffers as taught by *Meisel et al.* for a purpose of providing for analysis and collection of speech data for speech recognition to obtain optimal parameters during preprocessing.

A restriction requirement was made in the parent application, Application Serial No. 10/199,395, but the current claims of the application do not maintain the line of patentable distinctiveness. The current claims merely represent claims elected in the parent application, Application Serial No. 10/199,395.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1 to 3 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Barclay et al.* in view of *Meisel et al.*

Concerning independent claim 1, *Barclay et al.* discloses a speech recognition system, comprising:

"two or more clients, each client comprising the capability to receive audio speech from a user, store the audio speech in one or more buffers [in a raw uncompressed audio format], each buffer comprising a portion of the received audio speech, encode a buffer of the received audio speech before all of the audio speech is received, package the encoded buffer to receive audio speech into one or more packets to be transmitted over the internet before all of the audio speech is received, and transmit a packet of encoded audio speech over the internet before all of the audio speech is received" – a connection between the client and the server can be any communication channel including the Internet; a client includes a microphone 10 for accepting audio input ("capability to receive audio speech from a user") (column 4, line 62 to column 5, line 11: Figure 1); quantized feature data is delivered to dispatcher 26 where it may be temporarily buffered ("store the audio speech in one or more buffers, each buffer comprising a portion of the received audio speech") (column 5, lines 36 to

64); front-end 12 is a program for collecting the digitized speech, extracting a set of features, and quantizing those features ("encode a buffer of the received audio speech") (column 5, lines 4 to 11); such buffering does not detract from the real-time aspect since the buffering is to accommodate timing delays and synchronization that may be needed; the front end streams the quantized data to the dispatcher; "stream" is defined to send substantially continuously the data in real-time ("before all of the audio speech is received") (column 5, lines 48 to 55; column 7, lines 13 to 20); a client sends quantized speech data as message packets ("package the encoded buffer of received audio speech into one or more packets") (column 7, lines 48 to 59); the packets can be forwarded by the client dispatcher to the server ("transmit a packet of encoded audio speech") (column 7, lines 48 to 59); implicitly, a client/server architecture includes a plurality of clients ("two or more clients") serviced by a server;

"a server, said server comprising the capability to receive packets of encoded audio speech from at least two clients, decode each of the packets of audio speech and store the resultant raw speech into one or more buffers for the respective client, and evaluate the resultant raw speech received from each of the at least two clients" – server side 4 includes dispatcher 18 that physically receives the quantized features ("receive packets of encoded audio speech") (column 5, lines 21 to 35: Figure 1); server receives quantized speech data as message packets ("to receive packets of encoded audio speech") (column 7, lines 48 to 59); server dispatcher accepts and buffers messages (digitized quantized features) 60 before the recognizer is ready to receive and process the messages ("decode each of the packets of audio speech and store

resultant raw speech into one or more buffers for the respective client") (column 7, lines 26 to 40); speech recognizer/decoder 20 recognizes words from the quantized speech features ("evaluate the resultant raw speech received") (column 5, lines 21 to 35); implicitly, a client/server architecture enables a server to simultaneously service and buffer speech from a plurality of clients ("from each of the at least two clients simultaneously").

Concerning independent claim 1, the only element omitted by *Barclay et al.* is that clients store audio speech in buffers "in a raw uncompressed audio format". *Barclay et al.* discloses that a client buffers digitized speech parameters before it is sent, but omits buffering analog speech as it is received. However, it is well known to buffer data communications both upon reception and before transmission to permit processing. *Meisel et al.* teaches preprocessing for speech recognition, where a general approach is disclosed of a means for buffering raw analog or digitized speech data for analysis by collecting and storing the raw data. Optimal parameters can then be extracted by analysis. (Column 5, Lines 6 to 12) Thus, *Meisel et al.* suggests storing raw analog speech upon reception. It would have been obvious to one having ordinary skill in the art to include a feature of storing raw uncompressed audio speech in buffers as taught by *Meisel et al.* in a client/server speech processor/recognizer of *Barclay et al.* for a purpose of providing for analysis and collection of speech data for speech recognition to obtain optimal parameters during preprocessing.

Concerning claims 2 and 3, *Barclay et al.* discloses that a transcription or text is determined by the speech recognizer ("a result of the server's evaluation of the resultant raw speech received from the client"), the transcription is returned to the dispatcher, and the dispatcher returns the text to the client ("transmit a response to a client"); alternatively, the application program receives and understands the request from the client and performs the desired function (column 6, lines 5 to 25); a client has a browser 78 for displaying HTML (column 8, lines 36 to 47: Figure 4), which implicitly involves "a display screen".

Claims 5 and 6 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Barclay et al.* in view of *Meisel et al.* as applied to claim 1 above, and further in view of *Moshfeghi et al.*

Concerning claim 5, *Barclay et al.* discloses a server may provide web pages in HTML (column 8, lines 36 to 47), but does not expressly disclose two or more stored text formats, where the server selects a stored text format to a client as a result of the server's evaluation of the speech data. However, *Moshfeghi et al.* teaches an interactive voice response (IVR) system in a client/server architecture, where a server transmits Hypertext Markup Language (HTML) pages in accordance with personalization information stored in a user specific data store. The inclusion of such capability information allows the server to limit the use of or number of pixels in graphical objects in the HTML pages when the display is low resolution (or text only), or the bandwidth is limited so as produce unacceptably long download times. The HTML

page format is determined at user logon when the user speaks his ID and password ("as a result of the server's evaluation of the resultant raw speech received from the client"). (Column 4, Lines 37 to 65: Figure 1) It would have been obvious to one having ordinary skill in the art for a server to select a stored text format for a client as suggested by *Moshfeghi et al.* in the browser of *Barclay et al.* for the purpose of accommodating low resolution displays and low bandwidth network connections.

Concerning claim 6, *Barclay et al.* discloses client and server exchange information as message packets (column 7, lines 48 to 59); the transfer of information may be organized in conformance with the TCP/IP protocols (column 5, lines 16 to 20); implicitly, an internet operating in accordance with a TCP/IP protocol also partitions text into message packets ("the capability to partition a stored text format file into one or more packets") (column 7, lines 48 to 59).

Claims 8 to 13 and 18 to 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Barclay et al.* in view of *Meisel et al.*, and further in view of *Osborne et al.*

Concerning independent claims 9 and 11, the only elements omitted by *Barclay et al.* are buffers "organized as a linked list" and clients to "write the stored audio speech from a first buffer in a first set of buffers to a second buffer in a second set of one or more buffers", where the speech is then packaged and transmitted over the internet from the second buffer. However, *Osborne et al.* teaches a network interface, where it is stated that linked lists of buffers are typically used for transmitting from a transmit side

to a receive side. (Column 3, Lines 59 to 65) In one embodiment, a receive side includes a free buffer ring queue 56 ("a first set of one or more buffers") and buffers 64 and 66 ("a second buffer in a second set of one or more buffers"), where frame data received from a network interface is read from free buffer ring queue 56 to either buffer 64 or buffer 66. (Column 10, Lines 20 to 44: Figure 1B) Buffers are organized as linked lists. (Column 18, Line 58 to Column 19, Line 14: Figure 8) An objective is to ensure low overhead and prevent blocking of transmission of frames for other connections. (Column 3, Lines 10 to 58) It would have been obvious to one having ordinary skill in the art to organize buffers as a linked list and write from first to second buffers as taught by *Osborne et al.* in the browser of *Barclay et al.* for the purpose of ensuring low overhead and preventing blocking of transmission frames from other connections.

Concerning claims 8, 18, and 19, *Osborne et al.* teaches a network interface, where it is stated that linked lists of buffers are typically used for transmitting from a transmit side to a receive side (column 3, lines 59 to 65); buffers are organized as linked lists (column 18, line 58 to column 19, line 14: Figure 8).

Concerning claims 10 and 21, *Barclay et al.* discloses digitized speech is encoded as cepstra (column 5, lines 2 to 10), which is a compressed format for speech.

Concerning claims 12 and 13, *Barclay et al.* discloses that a transcription or text is determined by the speech recognizer ("a result of the server's evaluation of the resultant raw speech received from the client"), the transcription is returned to the dispatcher, and the dispatcher returns the text to the client ("transmit a response to a client"); alternatively, the application program receives and understands the request

from the client and performs the desired function (column 6, lines 5 to 25); a client has a browser 78 for displaying HTML (column 8, lines 36 to 47: Figure 4), which implicitly involves "a display screen".

Concerning claim 20, *Osborne et al.* teaches a receive side includes a free buffer ring queue 56 ("a first set of one or more buffers") and buffers 64 and 66 ("a second buffer in a second set of one or more buffers"), where frame data received from a network interface is read from free buffer ring queue 56 to either buffer 64 or buffer 66 (column 10, lines 20 to 44: Figure 1B); implicitly, there are "a predefined number" of second buffers, *i.e.* there are two buffers 64 and 66.

Claims 14 to 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Barclay et al.* in view of *Meisel et al.*, and further in view of *Osborne et al.* and *Moshfeghi et al.*

Concerning claims 14 and 17, *Barclay et al.* discloses a speech-enabled browser for displaying data via a graphical user interface (GUI), where the GUI may include a LISTEN button (column 8, lines 48 to 64), but does not expressly disclose a text-to-speech engine for converting a text format response to audio data, which is played through a speaker. However, text-to-speech engines are generally well known in interactive voice response (IVR) systems to provide two-way audio interaction without requiring a display screen. *Moshfeghi et al.* teaches an interactive voice response (IVR) system in a client/server architecture, where a client provides a text-to-speech synthesizer incorporated into a browser/Java® applet for audible messages to avoid

visual distraction to the user and to minimize storage requirements. (Column 1, Lines 39 to 45; Column 3, Lines 37 to 51; Column 4, Lines 25 to 36: Figure 1) It would have been obvious to one having ordinary skill in the art to incorporate a text-to-speech synthesizer into the browser of *Barclay et al.* as suggested by *Moshfeghi et al.* for the purpose of avoiding visual distraction to the user.

Concerning claim 15, *Barclay et al.* discloses a server may provide web pages in HTML (column 8, lines 36 to 47), but does not expressly disclose two or more stored text formats, where the server selects a stored text format to a client as a result of the server's evaluation of the speech data. However, *Moshfeghi et al.* teaches an interactive voice response (IVR) system in a client/server architecture, where a server transmits Hypertext Markup Language (HTML) pages in accordance with personalization information stored in user specific data store. The inclusion of such capability information allows the server to limit the use of or number of pixels in graphical objects in the HTML pages when the display is low resolution (or text only), or the bandwidth is limited so as produce unacceptably long download times. The HTML page format is determined at user logon when the user speaks his ID and password ("as a result of the server's evaluation of the resultant raw speech received from the client"). (Column 4, Lines 37 to 65: Figure 1) It would have been obvious to one having ordinary skill in the art for a server to select a stored text format for a client as suggested by *Moshfeghi et al.* in the browser of *Barclay et al.* for the purpose of accommodating low resolution displays and low bandwidth network connections.

Concerning claim 16, *Barclay et al.* discloses client and server exchange information as message packets (column 7, lines 48 to 59); the transfer of information may be organized in conformance with the TCP/IP protocols (column 5, lines 16 to 20); implicitly, an internet operating in accordance with a TCP/IP protocol also partitions text into message packets ("the capability to partition a stored text format file into one or more packets") (column 7, lines 48 to 59).

Response to Arguments

Applicant's arguments filed 13 March 2006 have been considered but are moot in view of the new grounds of rejection.

Conclusion

The prior art made of record and not relied upon is considered pertinent to Applicant's disclosure.

Naitoh discloses a speech encoder, where an analog speech signal input is applied via a buffer prior to encoding. (Column 1, Lines 44 to 48; Column 2, Lines 20 to 34)

Applicant's amendment necessitated the new grounds of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

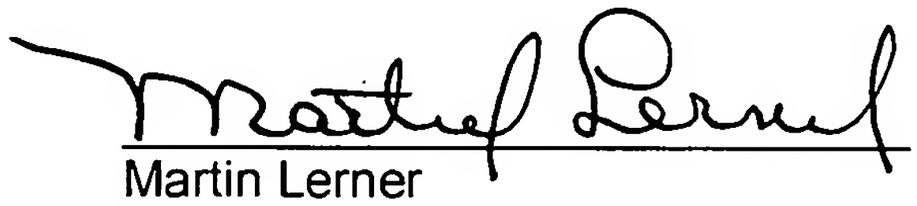
A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

ML
6/8/06



Martin Lerner
Examiner
Group Art Unit 2626